DOCKET SECTION

BEFORE THE
POSTAL RATE COMMISSION
WASHINGTON, D.C. 20268-0001

FEB 6 3-43 PM '98

POSTAL RATE AND FEE CHANGES, 1997

Docket No. R97-1

RESPONSES OF DOW JONES & COMPANY, INC. WITNESS SHEW TO INTERROGATORIES OF THE UNITED STATES POSTAL SERVICE (USPS/DJ-T1-1-6) (February 6, 1998)

Pursuant to the Commission's Special Rules of Practice, Dow Jones & Company, Inc. ("Dow Jones") hereby replies to the interrogatories posed by the United States Postal Service (USPS/DJ-T1-1-6) on January 23, 1998.

Each interrogatory is stated verbatim and is followed by the response.

Respectfully submitted,

Michael F. McBride

Samuel Behrends IV

Brenda Durham

Joseph H. Fagan

LeBoeuf, Lamb, Greene & MacRae, L.L.P.

1875 Connecticut Avenue, N.W., Suite 1200

Washington, DC 20009-5728

(202) 986-8000 (Telephone)

(202) 986-8102 (Facsimile)

February 6, 1998

Attorneys for Dow Jones & Company, Inc.

USPS/DJ-T1-1. Please assume that it is known a priori that an item type always contains one particular subclass of mail. Also assume that some of the observations of this item type in IOCS consist of "counted" (i.e., direct) observations and "uncounted" (i.e. mixed) observations, and that the a priori information is not imposed in the tally editing process.

- (a) In this case, would the subclass distribution of the counted observations for this item type accurately predict the subclass distribution of the uncounted observations for the same item type? If your answer is negative, please explain fully.
- (b) In this case, would the subclass distribution of the counted observations for this item type predict the subclass distribution of the uncounted observations for the same item type more accurately than a system-wide aggregate distribution of direct costs would? If your answer is negative, please explain fully.

RESPONSE:

The assumption on which these questions hinge is ambiguous. Had the assumption been more precise, I assume it would have read "Assume that it is known a priori that an item type always contains one particular subclass of mail and only that subclass of mail." The assumption can be illustrated by a simple, hypothetical example. Suppose there are only two item types, red sacks and blue sacks. Then, by the assumption above, red sacks contain exclusively (say) first-class letters and blue sacks contain exclusively (say) regular second class mail.

In that hypothetical circumstance, it would be trivially true that (a) the "subclass distribution for the item type" would perfectly predict the subclass distribution for uncounted items of the same item type (translation: if it the sack is blue, the mail must be regular second class), and therefore (b) no alternative predictor could do better. But that is nothing more than a convoluted way of repeating the assumption, namely that red sacks contain only first-class letters and blue sacks contain only regular second class mail.

Notice that if that assumption were true, there would be no mixed mail to worry about in the first place, since each item type would contain only one subclass. In other words, in the hypothetical circumstances in which I am being asked to evaluate a predictor of mixed mail costs, there would be no mixed mail.

USPS/DJ-T1-2. Please consider a mixed-mail IOCS tally that appears in the BCS cost pool

- (a) Please confirm that the mail handled in the BCS cost pool consists almost exclusively of letter-shape pieces that are compatible with letter automation equipment. If you do not confirm, please explain.
- (b) Please confirm that the mail handled in the FSM cost pool consists almost exclusively of machinable flat-shape pieces. If you do not confirm, please explain.
- (c) Would you expect the mail handled in the BCS cost pool to have a different subclass distribution than the mail handled in the FSM cost pool? If not, please explain.
- (d) Is it your testimony that observations of mail handlings in the FSM cost pool provide useful information for constructing subclass distributions for mixed-mail observations in the BCS cost pool? If not, please explain your testimony at page 22 (especially lines 16-18).
- (e) Does your testimony at page 22 (especially lines 16-18) imply that, in general, mixed-mail observations in letter cost pools would be distributed to subclass more accurately using information from both letter and non-letter cost pools than with information from letter cost pools alone? If not, please explain your testimony.
- (f) Does your testimony at page 22 (especially lines 16-18) imply that, in general, mixed-mail observations in letter automation cost pools would be distributed to subclass more accurately using information from both letter automation and manual cost pools than with information from letter automation cost pools alone? If not, please explain your testimony.

 RESPONSE:
- (a) (b) These questions concerning characteristics of mail in particular cost pools fall well outside the scope of my testimony. I did not discuss, nor are my conclusions predicated on, any particular stylized facts about characteristics of mail in particular cost pools. The characteristic of cost pools relevant to this proceeding is service cost responsibility, and little seems to be known about the services responsible for the majority of time spent by clerks and mail handlers.

- (c) The probability that any two distributions will be identical is normally very small, so I would not expect mail subclasses found in the BCS and FSM cost pools or, for that matter, any other pair of cost pools to be identically distributed.
- (d) (f) These questions refer to page 22, lines 16-18 of my testimony, where I say "The CPP assumption implies that activities in other cost pools provide no useful information on the services responsible for mixed mail and staff not handling mail in a cost pool. But that seems somewhat implausible."

My point was simple. Mr. Degen's assumptions treat each cost pool as if it were hermetically sealed, unrelated to anything else happening in the same facility, so the assumptions used to distribute a pool's cost of mixed mail and staff not handling mail are entirely inward-looking, and make no use of information from other cost pools. Yet it strains credulity to suppose that what is going on in those other cost pools would never cast any light on the services responsible for mixed mail and staff not handling mail in the pool being scrutinized.

Consider, for example, the "backup" explanation offered for staff not handling mail in manual sorting activities. According to that interpretation, manual sorting operations are scaled to provide the reserve capacity to handle peak loads of mail normally sorted automatically. In that event, the services to which staff not handling mail in manual operations are attributed should reflect in part the subclass distribution of mail in automated operations.

Whether or not one accepts the "backup" explanation, the general point remains. Cost pools are not in fact hermetically sealed off from one another, and it would be surprising if what is going on in one pool never could cast light on the appropriate service attribution of mixed mail and staff not handling mail in any other cost pool.

To say that some cost pools seem likely to exhibit connections to others is not to say that every cost pool necessarily casts light on the services responsible for costs in every other pool. Thus, the short answer to (d) - (f) is no, not necessarily. But for Mr. Degen's CPP assumption to be correct, activities in other cost pools could never provide any useful information about the services responsible for mixed mail and staff not handling mail in a cost pool. That is an extreme assumption, and one that strikes me as implausible.

USPS/DJ-T1-3. If the costs of not-handling mail activities in a cost pool were shown to be causally related to the volume of mail handled in the same cost pool, could it be reasonable to distribute such costs within the cost pool? Please explain.

RESPONSE:

To provide a sound basis for rate-making, costs should be distributed across services in the way that reflects cost causality. That is equally true whether the service volume that causes a cost is in some sense observed within the cost pool or in some other cost pool. In either case, the ultimate objective is to determine which services are responsible for what costs.

In investigating the cause of a pool's cost, whether one ought to confine attention to those service volumes observed within the pool or take a more complete view of service volumes at the facility is essentially an empirical matter of which approach will produce the most accurate estimates of service costs.

USPS/DJ-T1-4. Please refer to your testimony at page 28, lines 5-6.

- (a) Please confirm that "distributions of documented [i.e., direct] cost" (page 28, line 5) should read "distributions of mixed-mail cost." If you do not confirm, please explain why it is necessary to infer a direct cost distribution.
- (b) Please confirm that your statement, and the statement in MPA-T-2 to which you refer, are based on data provided in USPS-LR-H-305. If you do not confirm, please explain the basis for your statement.
- (c) What proportion of mixed-mail costs are distributed on five or fewer tallies? What proportion of total mail processing costs does this represent? Please provide the calculations to support your answer in electronic spreadsheet format.

RESPONSE:

(a) No. I did mean to refer to direct costs. The explanation follows.

Mr. Degen does not know the service responsibility for uncounted mixed-mail items (or, for that matter, staff not handling mail), so his methodology generally assumes it is identical in percentage terms to the service responsibility for the direct costs of items of the same type in the

cost pool. Thus, a service accounting for (say) 20% of the direct costs for a particular cost pool and item type (e.g., LSM sorting, flat trays) is assumed to be also responsible for 20% of the costs of uncounted mixed mail attributed to that pool and item type.

To apply that assumption, however, it is necessary to determine the distribution of direct costs for items of each particular type in the cost pool. That is not known, since the activities of clerks and mail handlers are not continuously monitored. The IOCS system provides sample information. It is necessary to infer from the sample data the direct cost distribution of services (i.e., the population distribution of direct costs). But a sample is almost never perfectly representative of the population, and statistical inferences from small samples can be prone to large errors.

In short, then, the problem is this. In simple, abstract terms, there is a cost variable X (mixed mail) whose service responsibility Mr. Degen does not know. He nevertheless wants to assign X to individual services, so he decides to assume that X has the same distribution of service responsibility as another cost, Y (direct cost). But he doesn't know the distribution of Y either; he has only sample information, and the samples are often too small to provide statistically reliable inferences. Thus, his methodology could go wrong not only because of his assumption that X is distributed the same as Y, but also because his estimate of the distribution of Y contains large errors.

- (b) In saying that hundreds of distributions of direct costs must be inferred from samples containing fewer than five direct tallies, I was relying on the direct testimony of witness Cohen cited in my testimony. Her testimony, I gather, was based on data provided in USPS-LR-H-305.
- (c) I do not know the proportion of mixed-mail costs distributed on the basis of five or fewer tallies, a proportion that in any event by itself would not be terribly informative. The key question is the cumulative magnitude of errors introduced into the analysis by the small samples upon which Mr. Degen's methodology relies. That cumulative error includes the errors from samples that, although larger than five, are still troublingly small from the vantage point of statistical reliability.

To determine the cumulative errors, it would be necessary to examine, for each service, the sampling errors introduced by samples of one, two, three and so forth, to the largest samples employed by Mr. Degen's methodology to distribute costs to the service. Had those calculations been made, one could quantify the sampling errors likely to be contained in each cost estimate. Mr. Degen presumably was aware that his assumptions required relying on samples whose small size could pose statistical problems, and I would have expected his analysis to address that topic and provide some quantification of the probable range of errors from that particular feature of his methodology.

USPS/DJ-T1-5. Please refer to your testimony at page 28, lines 7-8.

- (a) Please confirm that any costing system based on a statistical sample of mail processing activities over the course of a year would generally assign different costs to the same service in successive years. If you do not confirm, please explain.
- (b) Please confirm that the Postal Service's operations are not identical in every respect year after year. If you do not confirm, please explain.

RESPONSE:

(a), (b) Of course cost estimates from samples will vary with the samples themselves, and although I am not sure I know it from personal observation, it seems a safe bet that the Postal Service's operations are not identical in every respect year after year. But neither of these observations in any way affects the concern expressed in the paragraph that is the subject of this USPS interrogatory. For my concern was not that there would simply be some variability in cost assignments as a result of the small samples that Mr. Degen's methodology uses, but rather that there could be substantial random variability. In other words, the issue of degree is key here. I tried to make that clear by saying, in the cited paragraph,

This means that <u>substantial</u> elements of Mr. Degen's attributions of service cost are random. To understand the significance of that randomness, suppose that the Postal Service's operations were identical in every respect year after year, so no changes in service costs or volumes occurred. Mr. Degen's methodology would nevertheless be capable of attributing in successive years <u>quite different</u> costs to the same service. Those

random swings would reflect the <u>large</u> uncertainty associated with the small samples whose use is compelled by his methodology. (page 28, lines 7-13; emphasis added)

In short, a random error on the order of magnitude of 1% is not very worrying. But the reliance of Mr. Degen's methodology on many small samples raises the possibility of quite substantial random errors and, correspondingly, erratic swings in estimated service costs from year to year. Had Mr. Degen made the type of calculation that is sketched in my response to part (c) of USPS/DJ-T1-4 above, the magnitude of this problem would be clear.

USPS/DJ-T1-6. Please refer to your testimony at page 29.

- (a) Is it your testimony that you are not able to determine whether or not there is a better set of cost distribution assumptions than witness Degen's, because you are not sufficiently knowledgeable about Postal Service operations? If not, please explain your testimony at page 29, lines 4-7.
- (b) If you are not sufficiently knowledgeable about Postal Service operations to weigh the merits of various cost distribution assumptions, on what experience do you base your evaluation of witness Degen's methodology?

RESPONSE:

- (a) No. The fundamental impediment to evaluating Mr. Degen's methodology is that it consists of extensive assumptions that he has apparently not taken the trouble to test. And although there is no reason that data to test his assumptions could not be obtained from within the Postal Service, I am not aware of publicly available information that would permit any comprehensive testing of his assumptions.
- (b) I believe what page 29 of my testimony says is that (1) if assumptions are to be used, I happily leave the task of evaluating competing assumptions to others more knowledgeable about Postal Service operations, because (2) information in the hands of those having long

experience with Postal System operations and cost data should be able to cast light on the plausibility of Mr. Degen's assumptions and on whether other procedures for dealing with the cost of mixed mail and staff not handling mail might produce better estimates of service costs.

As for the experience that I bring to evaluating the cost study of Mr. Degen, I suppose it includes lessons I have learned from analyzing cost and price issues in a number of regulated industries. Those lessons include the dysfunctionality of Fully Distributed Cost studies, the tendency of regulated organizations to exaggerate the difficulties of obtaining data to do a meaningful cost study, and the regularity with which assumptions about service costs turn out to be badly wrong.

DECLARATION

I, William B. Shew, declare under penalty of perjury that the foregoing answers are true and correct, to the best of my knowledge, information and belief.

Date: 2 6 98

CERTIFICATE OF SERVICE

I hereby certify that I have this day served the foregoing document upon all participants of record in this proceeding in accordance with section 12 of the Rules of Practice.

Brenda Durham

Dated: February 6, 1998